

## SEQUENCE LISTING

&lt;110&gt; Valtion teknillinen tutkimuskeskus

&lt;120&gt; Improved biotechnical production method

&lt;130&gt; vtt96

&lt;140&gt; 19991781

&lt;141&gt; 1999-08-20

&lt;160&gt; 8

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 2868

&lt;212&gt; DNA

&lt;213&gt; Trichoderma reesei

&lt;220&gt;

&lt;221&gt; gene

&lt;222&gt; (1523)..(1950)

&lt;223&gt; hfb1

&lt;400&gt; 1

|            |            |             |            |             |             |      |
|------------|------------|-------------|------------|-------------|-------------|------|
| tttgtatggc | tggatctcga | aaggcccttg  | tcatecgcaa | gcgtgggctaa | tatcgaatga  | 60   |
| gggacaccga | gttgcatatc | tcctgatcat  | tcaaacgaca | agtgtgaggt  | aggcaatcct  | 120  |
| cgtatcccat | tgctgggctg | aaagcttcac  | acgtatcgca | taagcgtctc  | caaccagtgc  | 180  |
| ttaggtgacc | cttaaggata | cttacagtaa  | gactgtatta | agtcagtcac  | tctttcactc  | 240  |
| gggctttgaa | tacgatcctc | aatactcccg  | ataacagtaa | gaggatgata  | cagcctgcag  | 300  |
| ttggcaaatg | taagcgtaat | taaactcagc  | tgaacggccc | ttgttgaaag  | tctctctcga  | 360  |
| tcaaagcaaa | gctatccaca | gacaaggggt  | aagcaggctc | actcttccta  | cgccttggat  | 420  |
| atgcagcttg | gccagcatcg | cgcattggcca | atgatgcacc | cttcacggcc  | caacggatct  | 480  |
| cccgttaaac | tcccctgtaa | cttggcatca  | ctcatctgtg | atcccaacag  | actgagttgg  | 540  |
| gggctgcggc | tggcggatgt | cggagcaaag  | gatcacttca | agagcccaga  | tccggttggg  | 600  |
| ccattggcaa | tggatctaga | ttcggcacct  | tgatctcgat | caactgagaca | tgggtgagttg | 660  |
| cccgagcgca | ccacaactcc | ccctgtgtca  | ttgagtcctc | atatgcgtct  | tctcagcgtg  | 720  |
| caactctgag | acggattagt | cctcacgatg  | aaattaactt | ccagcttaag  | ttcgtagcct  | 780  |
| tgaatgagtg | aagaaatttc | aaaaacaaac  | tgagtagagg | tcttgagcag  | ctggggtggg  | 840  |
| acgcccctcc | tcgactcttg | ggacatcgta  | cggcagagaa | tcaacggatt  | cacacctttg  | 900  |
| ggtcgagatg | agctgatctc | gacagatacg  | tgcttcacca | cagctgcagc  | tacctttgcc  | 960  |
| caaccattgc | gttccaggat | cttgatctac  | atcacccgag | cacccgagcc  | aggacggaga  | 1020 |
| gaacaatccg | gccacagagc | agcaccgcct  | tccaactctg | ctcctggcaa  | cgtcacacaa  | 1080 |
| cctgatatta | gatatccacc | tgggtgattg  | ccattgcaga | gaggtggcag  | ttgggtgatac | 1140 |
| cgactggcca | tgcaagacgc | ggccgggcta  | gctgaaatgt | ccccgagagg  | acaattggga  | 1200 |
| gcgtctatga | cggcgtggag | acgacgggaa  | aggactcagc | cgctcatgtt  | tggttgccaat | 1260 |
| ttgagattgt | tgaccgggaa | agggggggacg | aagaggatgg | ctgggtgagg  | tggtattggg  | 1320 |
| aggatgcatc | attcgactca | gtgagcgatg  | tagagctcca | agaatataaa  | tatcccttct  | 1380 |
| ctgtcttctc | aaaatctcct | tccatcttgt  | ccttcacagc | caccagagcc  | agcctgaaca  | 1440 |
| cctccagtca | acttccctta | ccagtacatc  | tgaatcaaca | tccattcttt  | gaaatctcac  | 1500 |
| cacaaccacc | atcttcttca | aaatgaagtt  | cttcgccatc | gccgctctct  | ttgccgccgc  | 1560 |
| tgccgttgcc | cagcctctcg | aggaccgcag  | caacggcaac | ggcaatgttt  | gccctcccgg  | 1620 |
| cctcttcagc | aacccccagt | gctgtgccac  | ccaagtcctt | ggcctcatcg  | gccttgactg  | 1680 |
| caaagtccgt | aagttgagcc | ataacataag  | aatcctcttg | acggaaatat  | gccttctcac  | 1740 |
| tcctttacc  | ctgaacagcc | tcccagaacg  | tttacagcgg | caccgacttc  | cgcaacgtct  | 1800 |
| gcgccaaaac | cggcgcccag | cctctctgct  | gcgtggcccc | cggttgtaagt | tgatgcccc   | 1860 |
| gctcaagctc | cagtctttgg | caaaccatt   | ctgacaccca | gactgcaggc  | cggccagcct  | 1920 |
| cttctgtgcc | agaccgccgt | cgggtgcttg  | gatgcccgcc | cggggtcaag  | gtgtgcccg   | 1980 |
| gagaaagccc | acaaagtgtt | gatgaggacc  | atttccggta | ctgggaaagt  | tggctccacg  | 2040 |
| tgtttgggca | ggtttgggca | agttgtgtag  | atattccatt | cgtacgccat  | tcttattctc  | 2100 |
| caatatttca | gtacactttt | cttcataaat  | caaaaagact | gctattctct  | ttgtgacatg  | 2160 |
| cgggaaggga | acaattgctc | ttggtctctg  | ttatttgcaa | gtaggagtg   | gagattcgcc  | 2220 |

|             |             |             |             |            |            |      |
|-------------|-------------|-------------|-------------|------------|------------|------|
| ttagagaaaag | tagagaagct  | gtgcttgacc  | gtggtgtgac  | tcgacgagga | tggactgaga | 2280 |
| gtgttaggat  | taggtcgaac  | ggtgaagtgt  | atacaggatc  | gtctggcaac | ccacggatcc | 2340 |
| tatgacttga  | tgcaatgggtg | aagatgaatg  | acagtgttaag | aggaaaagga | aatgtccgcc | 2400 |
| ttcagctgat  | atccacgcca  | atgatacagc  | gatatacctc  | caatatctgt | gggaacgaga | 2460 |
| catgacatat  | ttgtgggaac  | aacttcaaac  | agcgagccaa  | gacctcaata | tgcacatcca | 2520 |
| aagccaaaca  | ttggcaagac  | gagagacagt  | cacattgtcg  | tcgaaagatg | gcacgtacc  | 2580 |
| caaatcatca  | gctctcatta  | tcgcctaaac  | cacagattgt  | ttgccgtccc | ccaactccaa | 2640 |
| aacgttacta  | caaaagacat  | gggcgaatgc  | aaagacctga  | aagcaaacc  | tttttgcgac | 2700 |
| tcaattccct  | cctttgtcct  | cggaaatgatg | atccttcacc  | aagtaaaaga | aaaagaagat | 2760 |
| tgagataata  | catgaaaagc  | acaacggaaa  | cgaagaacc   | aggaaaagaa | taaatctatc | 2820 |
| acgcaccttg  | tccccacact  | aaaagcaaca  | gggggggtaa  | aatgaaat   |            | 2868 |

&lt;210&gt; 2

&lt;211&gt; 3585

&lt;212&gt; DNA

<213> *Trichoderma reesei*

&lt;220&gt;

&lt;221&gt; gene

&lt;222&gt; (1191)..(1593)

&lt;223&gt; hfb2

&lt;400&gt; 2

|             |            |             |             |             |             |      |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| ctcgagcagc  | tgaagcttgc | atgcctgcat  | cctttgtgag  | cgactgcac   | cattttgcac  | 60   |
| acactgccgt  | cgacgtctct | cttccgacct  | tggccagctg  | gacaagcaac  | acaccaatga  | 120  |
| cgctttgtat  | tattagagta | tatgcaagtc  | tcaggactat  | cgactcaact  | ctaccaccg   | 180  |
| aggacgatcg  | cggcacgata | cgccctcggt  | ctcattggcc  | caagcagacc  | aactgcccct  | 240  |
| ggagcaagat  | tcagcccaag | ggagatggac  | ggcagggcac  | gccaggcccc  | caccaccaag  | 300  |
| ccactccctt  | tggccaaatc | agcttgcattg | tcaagagaca  | tcgagctgtg  | ccttgaaatt  | 360  |
| actaacaacc  | agggatggga | aacgaagcct  | gcttttggaa  | agacaacaat  | gagagagaga  | 420  |
| gagagagggg  | gagagacaat | gagtgccaca  | aacctggtag  | tgctccgcca  | atgctgtctga | 480  |
| aatgtcacat  | ccgagtcctg | gggcctctgt  | gagaatgtcc  | agagtaatac  | gtgttttgcg  | 540  |
| aatagtcctc  | tttcttgagg | actggatacc  | tacgataccc  | tttttgagtt  | gatgcggtgc  | 600  |
| tttcgaagta  | ttatctggag | gatagaagac  | gtctaggtaa  | ctacacaaaa  | ggcctatact  | 660  |
| ttggggagta  | gccaacgaa  | aggtaactcc  | tacggcctct  | tagagccgtc  | atagatccta  | 720  |
| cagcctcttg  | gagccgtcat | agatcacatc  | tgtgtagacc  | gacattctat  | gaataatcat  | 780  |
| ctcatcatgg  | ccacatacta | ctacatacgt  | gtctctgcct  | acctgacatg  | tagcagtggc  | 840  |
| caagacacca  | aggccccagc | atcaagcctc  | cctacctatc  | ccttccattg  | tacagcggca  | 900  |
| gagagattgc  | gatgagccct | ctccctacct  | acagacggct  | gacaatgtcc  | gtataccacc  | 960  |
| agccaacgtg  | atgaaaacaa | ggacatgagg  | aacgcctgc   | gagagctgga  | agatgaagag  | 1020 |
| ggccagaaaa  | aaaagtataa | agaagacctc  | gattcccgc   | atccaacaat  | cctttccatc  | 1080 |
| ctcatcagca  | cactcatcta | caaccatcac  | cacattcact  | caactcctct  | ttctcaactc  | 1140 |
| tccaaacaca  | aacattcttt | ggtgaatacc  | aaccatcacc  | acctttcaag  | atgcagttct  | 1200 |
| tcgccgtcgc  | cctcttcgcc | accagcgccc  | tggctgctgt  | ctgccctacc  | ggcctcttct  | 1260 |
| ccaaccctct  | gtgctgtgcc | accaacgtcc  | tcgacctcat  | tggcggtgac  | tgcaagaccc  | 1320 |
| gtatgttgaa  | ttccaatctc | tgggcatcct  | gacattggac  | gatacagttg  | acttacacga  | 1380 |
| tgctttacag  | ctaccatcgc | cgtcgacact  | ggcgccatct  | tccaggctca  | ctgtgccagc  | 1440 |
| aagggtccca  | agcctctttg | ctgcgttgct  | cccgtggtaa  | gtagtgtctg  | caatggcaaa  | 1500 |
| gaagtaaaaa  | gacatttggg | cctgggatcg  | ctaactcttg  | atatcaaggc  | cgaccaggct  | 1560 |
| ctcctgtgcc  | agaaggccat | cggcaccttc  | taaagcaatg  | gcttgcttta  | ctgccggcag  | 1620 |
| tctttgagaa  | ctctgggctc | acaaaagacg  | acttgcattg  | atcatggggg  | ctcgcaaatg  | 1680 |
| ggaggatttg  | gaggggattg | aggctggggt  | tggcctatta  | gaggattgca  | taatggaaga  | 1740 |
| tttgcgagca  | ggacatagac | gtatctagag  | ttctagtcaa  | tacattattg  | aaaagtggga  | 1800 |
| gtatacctat  | cgctggcact | ggtatcttga  | agatatcttc  | tcttcttggtg | aggttatgta  | 1860 |
| tggcaatcag  | tcgaaatcta | tttgaagaca  | gagctcaagc  | ttcaaacatt  | cacctgngaa  | 1920 |
| ttgaccattt  | tgtttcgatg | gttgacgttg  | gtgggtgtca  | cttctgcaat  | catgtacgag  | 1980 |
| cacaagtata  | gcagtatctc | atctgatctg  | catctgggta  | aatgtcgcca  | ctctacctag  | 2040 |
| gtaccaataa  | aataccgaat | tggtcagctc  | tcgggtgaca  | aaccggcccc  | cttttcgacc  | 2100 |
| gtgctctgtc  | caattctagg | cttgtaaatg  | gttctctgact | gtgataaacc  | ttggagctan  | 2160 |
| cataactttac | cttacaataa | atccaactgc  | cggcacttgc  | ttcccttcac  | ccaaccactc  | 2220 |
| gcaaactatca | cgaacacctg | ctcgatcccc  | tgtccgaaat  | ctgcttggca  | acgtatcatc  | 2280 |
| acaaatcata  | cacacagaca | aaaaggagcc  | aaagcagcaa  | tggcaagaca  | ccgaggccgg  | 2340 |

```

cagcgcgccc gtcgccgttt ttaaaaagcg aagcgcgcaaag ggcaaagcca acctgcgcaa 2400
acgaacaacg aagccttccc cccgccgcga gcgacagcga cagcgacagc gacttttcct 2460
cgtcggaaga cgaagccggg cacagagtca agaggcgcaa gaggacggcc gtcgtcaccg 2520
ccgccgcgga ggggcgccgc gccagcaac cgggacgacg gcggcgggcg aacagccgcc 2580
ttcacggcca acagaagcgt cccgattgct gacagcaacg acgcgaccaa gcacagcaac 2640
tggtacgacg aggacgcaaa ggacgcgctc tcggcaaaga acctcctcgg atcttcgaga 2700
gcgtccaagg acgcgcagcc agacggcacg tacaagggcc tggcgaacca gacgtccttt 2760
atacaaaaga atccggatgc gccccggaag acagttgggc ccgtcaaggc gcctaccaac 2820
atccgcaccg tcaccattac agattatgcc ccggacacgt gtaaagagtg agtttgcatc 2880
aatagccaga atcccccccc ccgataccgt acattgagca tatgctgact cgtcataatc 2940
tttctagtta tcgcataacc ggctatataa gtactccctt tttccatgat tattccagtc 3000
gcgtactgac atttctagga gcctttactg tggttttggc gacaattgca agtatcttca 3060
cgcgagagaa gacctcaagg caggctggca gctggatcaa gagggggaaa aggtcaccaa 3120
gggcaagaag aacctggggg gaacggtagt ggccagcgcg aaccggaaca aggccaaggt 3180
ggacgagggc gacgacgacg acgacgaaga ggcatgctc gagaacattc cgtttgcctg 3240
catcatctgc agggaatcgt acaaggagcc gattgtgacg aggtgcgggc actacttttg 3300
cctgccgtgc gctctgcagc ggtacaagaa ggatccgacg tgtgcggcgt gtggctcggg 3360
cacgaatggc gtgtttaatt cggcgacgag gttgaagaag ctgctggaga agaagaggga 3420
gagggcggcc aggaggagac aggaggcgat agagaggggc gaggaagtca gtgatgaaga 3480
ggaggaggag gaggaggact gatgatgatg gggcnagatg acgatgcagg tcgactctag 3540
agatccccgg taccgagctc gaattcatcg atgatatcag atccc 3585

```

<210> 3  
 <211> 63  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: 5' primer

```

<400> 3
actacacgga ggagctcgac gacttcgagc agccccgagct gcacgcagag caacggcaac 60
ggc 63

```

<210> 4  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: 3' primer

```

<400> 4
tcgtacggat cctcaagcac cgacggcggt 30

```

<210> 5  
 <211> 13  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: linker

```

<400> 5
Pro Gly Ala Ser Thr Ser Thr Gly Met Gly Pro Gly Gly
  1              5              10

```

<210> 6  
 <211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: linker

<400> 6

Gly Thr Leu Val Pro Arg Gly Pro Ala Glu Val Asn Leu Val  
1 5 10

<210> 7

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: 5'primer

<400> 7

gaattcggtgta ccctcggtccc tcgcggtccc gccgaagtga acctggtg

48

<210> 8

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: 3'primer

<400> 8

tgaattccat atgctaaccc cgtttcatct ccag

34